AD-A121497

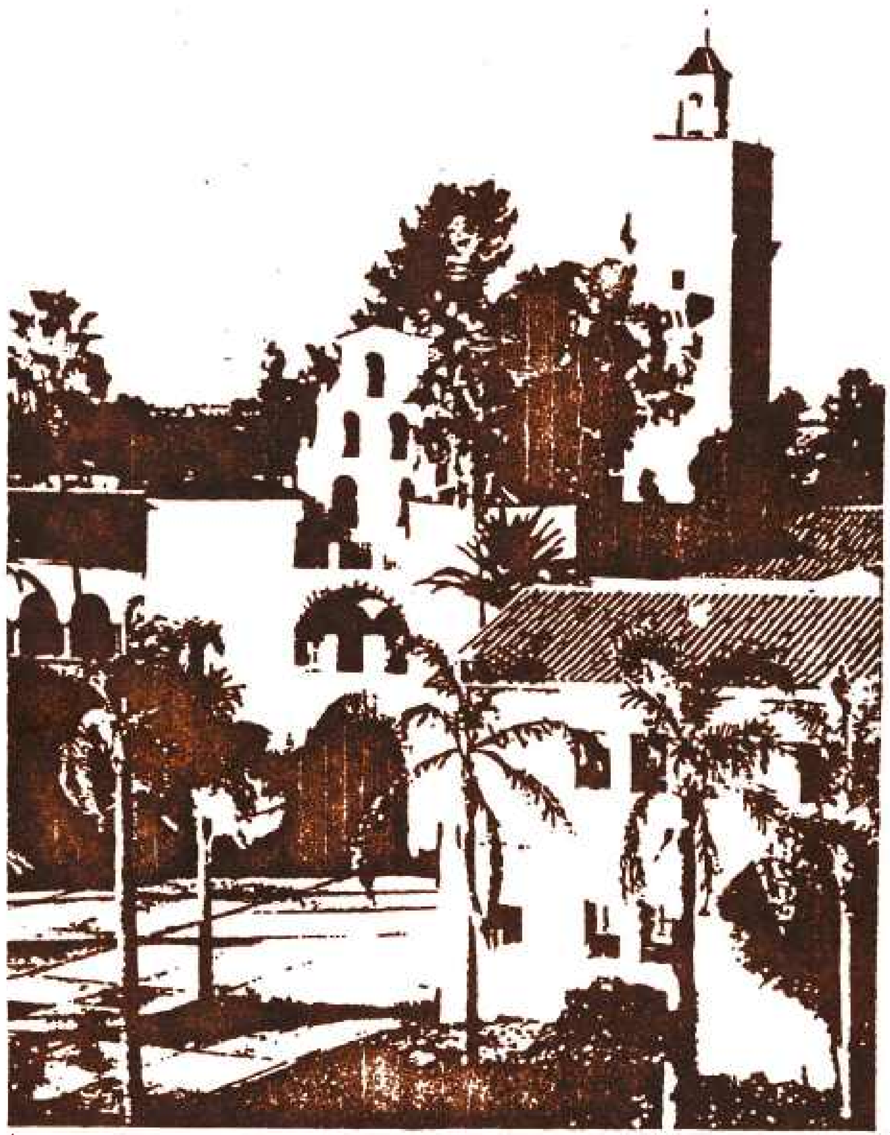
SDSU

2-82

SERIES IN STATISTICS AND BIOSTATISTICS

SIGNAL DETECTION FOR STOCHASTIC PROCESSES WITH STATIONARY INDEPENDENT SYMMETRIC INCREMENTS

C. B. BELL AND R. AHMAD



San Diego State University
Department of Mathematical Sciences

SIGNAL DETECTION FOR STOCHASTIC PROCESSES WITH STATIONARY INDEPENDENT SYMMETRIC INCREMENTS

C. B. Bell and R. Ahmad*

San Diego State University

University of Strathclyde, Glasgow*

1. Introduction and Summary

The object of this paper is to develop statistical signal detection techniques for a wide class of stochastic processes with stationary independent symmetric increments. That is, the basic premise in this situation is that the increments from a regularly-sampled continuous parameter process are i.i.d. with a continuous distribution satisfying: $F(x) + F(-x) = 1$ for all x. Taking symmetry about zero is no loss of generality, since one can initially take symmetry amount some constant c and then one can re-define the process to be symmetric about zero.

Let $\{Z(t), t \geq 0\}$ be a stochastic process of this type, then the realized data from signal detection viewpoint will be of the following two types:

(a) Historical Data: $\underset{\sim}{X} = (X_1, X_2, \ldots, X_m)$, where $X_r = Z(r\Delta) - Z((r-1)\Delta)$, $Z(0) = 0$ and $\Delta > 0$.

(b) Two-sample Data: $(\underset{\sim}{X}, \underset{\sim}{Y}) = (X_1, X_2, \ldots, X_m; Y_1, Y_2, \ldots, Y_n)$,

This work was principally supported by the Office of the Naval Research through Grant No. N00014-80-C-0208.

X_j's and Y_j's are defined similar to (a) above with respect to stochastic processes $\{Z(t), t \geq 0$ with law $\mathcal{L}_Z\}$ and $\{Z'(t), t \geq 0$ with law $\mathcal{L}_{Z'}\}$.

The family of all stochastic processes with stationary independent symmetric increments, in the sequel, will be denoted by $\Omega(\text{SISI})$. Three types of detection problems will be treated. These are outlined below in terms of pure noise (PN) and noise plus signal (N + S); see Bell (1964a).

(i) Goodness-of-fit Detectors. Here, we detect the problem of PN: $\mathcal{L} = \mathcal{L}_0 \in \Omega(\text{SISI})$ against N + S: $\mathcal{L} \neq \mathcal{L}_0$, where $\mathcal{L}_0$ is completely specified.

(ii) Class-fit-Detectors. This involves detecting PN: $\mathcal{L} \in \Omega(\text{SISI})$ versus N + S: $\mathcal{L} \notin \Omega(\text{SISI})$.

(iii) Two-sample Detectors. This problem is to detect PN: $\mathcal{L}_1 = \mathcal{L}_2$ ($\mathcal{L}_i \in \Omega(\text{SISI})$, $i = 1, 2$) against N + S: $\mathcal{L}_1 \neq \mathcal{L}_2$, where again $\mathcal{L}_1$ and $\mathcal{L}_2$ are in $\Omega(\text{SISI})$.

Let α denote the probability that the detector will produce a false alarm (PFA); and denote by β the probability that the detector will produce a false dismissal (PFD). The detectors proposed in this investigation are optimal in the sense that for a fixed PFA α, the procedure has minimum PFD, β. The organization of the paper is as follows.

In Section 2 some examples of laws from the family $\Omega(\text{SISI})$ are given. Section 3 starts with some basic terminology used. Also, discussion about minimal sufficient statistics (M-S-S), maximal statistical noise (M-S-N) and a non-parametric property for the underlying detection statistics is given. In part 4 relevant distributions, permutations and alternative statistics with some examples are discussed. Section 5 treats goodness-of-fit detectors. In the next section 6 class-fit detectors are developed. The two-sample detectors are dealt with in section 7. The final part concludes with point estimation and confidence bounds for F in Ω(SISI) which are quite useful in detection methodology.

For various structural results and properties in the case of stochastic processes belonging to the class Ω(SISI), the relevant references are Doob (1953), Feller (1966) and Basawa and Rao (1980). A different but related concept of symmetry is investigated by Bell and Haller (1969), Bell and Smith (1969, 1972), and Ahmad (1974). Wiener-Levy processes which are closely connected with processes of the class Ω(SISI) are treated in Bell et al. (1980). The detection techniques developed in this paper carry over to discrete time as well as continuous time parameter stochastic processes in the family Ω(SISI).

2. Some Examples of Laws from Ω (SISI)

Example 2.1. (Wiener). Let $\{Z(t): t \geq 0\}$ be a Gaussian process satisfying (a) $Z(0) = 0$; (b) $E(Z(t)) = 0$, (c) $\mathrm{Cov}\,(Z(s), Z(t)) = \sigma \min\,(s,t)$. Then for any $\Delta > 0$, $Y_1, \ldots, Y_n$ are i.i.d. $N(0, \sigma^2\Delta)$, when $Y_j = Z(j\Delta) - Z([j-1]\Delta)$. That is the process has SISI's, or $\mathcal{L} \in \Omega$(SISI).

Example 2.2. (Compound Poisson). Let $Y_1, Y_2, \ldots, Y_n, \ldots$ be i.i.d. $\mathcal{U}(-\theta,\theta)$; $\{N(t): t \geq 0\}$ be a H-P-P (Homogeneous Poisson Process) with parameter λ; and $Z(t) = \sum_1^{N(t)} Y_j$. Then $\{Z(t): t \geq 0\}$ is a Compound Poisson Process. Further, for each $\Delta > 0$, $Y_1, \ldots, Y_n$ are i.i.d. F^*, with $F^*(x) + F^*(-x) = 1$ for all x, when $Y_j = Z(j\Delta) - Z([j - 1]\Delta)$.

Example 2.3. (Random Walk). Let $W_1, W_2, \ldots, W_n, \ldots$ be i.i.d. double exponential, $D - E(\lambda)$, i.e., $f_W(x) = \frac{1}{2} e^{-\lambda|x|}$; and $Z(r) = \sum_1^r W_j$. Then $\{Z(r): r = 1, 2, \ldots\}$ has a law $\mathcal{L}$ in Ω(SISI).

Example 2.4. (Symmetric Stable Distributions). A distribution $F(x)$ is said to be strongly unimodal, if and only if, the convolution of G with any unimodal distribution is unimodal (normal and Wishart distributions are strongly unimodal). If F and G are symmetric about zero and unimodal distributions, then so is their convolution $F*G$. Furthermore, all stable distributions (as defined below) with characteristic functions given by $\exp(-|t|^a)$, $0 < a \leq 2$ are unimodal. Let $X, X_1, X_2, \ldots$ denote mutually independent random variables with a common distribution H and set $X_n^* = X_1 + X_2 + \ldots + X_n$. The distribution H is stable if for each n there exist constants $b_n > 0$ and c_n such that $X_n^* \overset{d}{=} b_n X + c_n$ and H is not concentrated at the origin. Consequently, if $X_1, X_2, \ldots$ are i.i.d. are stable and symmetric about zero, then $\{Z(t) = X_r^*,$ $r = 1, 2, \ldots\}$ has a law $\mathcal{L}$ in Ω(SISI).

Remark 2.1. Stable distributions are natural generatizations of the normal family. Only the norming constants $b_n = n^{1/a}$ are possible; a is called the characteristic exponent of the distribution $H(\cdot)$. All stable distributions are continuous. For many applications and other results for stable distributions see Feller (1966).

3. Sufficient Statistics and the Non-parametric Property

In developing the statistics to be used it is convenient to delineate two types of distribution-free-ness.

Definition 3.1. (a) A statistic $T(\cdot)$ is NPDF wrt a family Ω' of stochastic-process laws, if there exists a cpf $Q(\cdot)$ such that

$$P\{T(\underset{\sim}{X}) \leq t | \mathcal{L}\} = Q(t) \quad \text{for all} \quad t, \quad \text{and for all} \quad \mathcal{L} \in \Omega'.$$

(b) A family $\{T^*(\cdot;\mathcal{L}): \mathcal{L} \in \Omega'\}$ is PDF wrt a family Ω', if there exists a cpf Q^* such that

$$P\ T^*(\underset{\sim}{x};\mathcal{L}) \leq t|\mathcal{L}\} = Q^*(t) \quad \text{for all} \quad t \quad \text{and for all} \quad \mathcal{L} \in \Omega'.$$

Example 3.1. Consider Example 2.1 with $\Omega' = \{\text{WLP}: \ \sigma > 0\}$ and data $\underset{\sim}{Y} = (Y_1, \ldots, Y_9)$. Let $T(\underset{\sim}{Y}) = \frac{5}{4} [\sum_1^4 y_j^2] [\sum_5^9 y_j^2]^{-1}$ and $T^*(\underset{\sim}{Y}, \mathcal{L}) = \sigma^{-2} [\sum_1^9 y_j^2]\Delta^{-1}$. Then $T(\cdot)$ is NPDF wrt Ω', with $Q = F(4; 5)$, and $T^*(\cdot,\cdot)$ is PDF wrt Ω' with $Q^* = \chi_9^2$.

The interest here is in the wider family Ω (SISI). In order to construct the PDF and NPDF statistics, one needs the minimal sufficient statistic (M-S-S) and its complementary statistic, the maximal statistical noise, M-S-N, to be defined below.

Notation. Let $\underset{\sim}{X} = (X_1, \ldots, X_n)$ be the vector of increments, i.e., $X_j = Z(j\Delta) - Z([j-1]\Delta)$; let $S_1(\underset{\sim}{X}) = (|X|(1), \ldots, |X|(n))$, the vector of ordered absolute values; let $N_1(\underset{\sim}{X}) = (\underset{\sim}{\varepsilon}, \underset{\sim}{R}^*)$, where

$$\underset{\sim}{\varepsilon} = [\varepsilon(X_1), \varepsilon(X_2), \ldots, \varepsilon(X_n)], \text{ and } \underset{\sim}{R}^* = [R(|X_1|), \ldots, R(|X_n|)];$$

and $\delta_1(\underset{\sim}{X}) = [S_1(\underset{\sim}{X}), N_1(\underset{\sim}{X})]$.

Theorem 3.1. (a) $S_1(\underset{\sim}{X})$ is the M-S-S for Ω(SISI), and it is complete.

(b) $S_1(\underset{\sim}{X})$, $\underset{\sim}{\varepsilon}$ and $\underset{\sim}{R}^*$ are mutually independent

(c) $\delta_1(\cdot)$ is 1-1 a.e.

Definition 3.2. Let $S(\underset{\sim}{X})$ be a M-S-S for Ω'; $\delta(\underset{\sim}{X}) = [S(\underset{\sim}{X}), N(\underset{\sim}{X})]$ be 1-1 a.e., and $N(\underset{\sim}{X})$ be independent of $S(\underset{\sim}{X})$. Then

(a) $\delta(\cdot)$ is called the BDT (basic data transformation) for Ω'

and

(b) $N(\underset{\sim}{X})$ is called the M-S-N (maximal statistical noise) for Ω'.

Theorem 3.2. (a) $\delta_1(\underset{\sim}{X}) = [S_1(\underset{\sim}{X}), N_1(\underset{\sim}{X})]$ is the BDT for Ω(SISI);

(b) $N_1(\underset{\sim}{X}) = [\underset{\sim}{\varepsilon}, \underset{\sim}{R}^*]$ is M-S-N for Ω'.

The general rule for employing these statistics in signal dection is as follows.

Rule of Thumb. (A) For situations involving a specific law, (e.g., PN: $\mathcal{L} = \mathcal{L}_0$), employ a PDF statistic based on the M-S-S,

(B) For situations involving the underlying structure of the family e.g., PN: $\mathcal{L} \varepsilon \Omega'$ or PN: $\mathcal{L}_1 = \mathcal{L}_2$, employ an NPDF statistic based on the M-S-N.

This overall principle will be followed in the sequel. However, in order to make efficient use of this principle, one should consider several alternate versions of the M-S-S and the M-S-N.

4. Distributions, Permutations and Alternate Statistics.

The only versions of M-S-S and M-S-N, known to the authors, are those related to maximal invariants or permutation statistics. $S_1(\underset{\sim}{X})$, given above, is an appropriate maximal invariant, as is $N_1(\underset{\sim}{X})$.

The set of permutations of interest here is the Sign-Time group, S_n^*.

Definition 4.1. (a) $S_n^* = \{$all permutations of coordinates of $\underset{\sim}{X}$ and changes of signs of coordinates.$\}$

(b) $S_n^*(\underset{\sim}{X}) = \{\gamma(\underset{\sim}{X}): \gamma \varepsilon S_n^*\}$ is the S_n^* orbit of $\underset{\sim}{X}$.

Example 4.1. Let $n = 2$, and $\underset{\sim}{x} = (-5.6, 0.9)$ then, its orbit is

$$\begin{aligned} S_2^*(\underset{\sim}{x}) &= \{\gamma(\underset{\sim}{x}): \gamma \varepsilon S_2^*\} \\ &= \{(-5.6, 0.9), (-5.6, -0.9), (5.6, 0.9), (5.6, -0.9), \\ &\quad (0.9, -5.6), (-0.9, -5.6), (0.9, 5.6), (-0.9, 5.6)\}. \end{aligned}$$

Theorem 4.1. (a) S_n^* is a group of order $(n!)(2^n)$

(b) S_n^* is a wreath product group.

(c) The orbit, $S_n^*(\underset{\sim}{X})$ contains $(n!)(2^n)$ points for almost every $\underset{\sim}{X}$.

Based on these permutations, one can now give useful additional versions of the M-S-S and the M-S-N.

<u>Theorem 4.2.</u> $S_n^*(\underset{\sim}{X})$, the S_n^* orbit of Definition 4.1, is a M-S-S. In order to construct the M-S-N based on S_n^*, one needs the following definitions

<u>Definition 4.2.</u> (a) A (measurable, real-valued) function $h(\cdot)$ is called a B-Pitman function wrt a set, S', of permutations, and a family Ω', of stochastic-process laws if

$$P\{h(\underset{\sim}{X}) = h(\gamma(\underset{\sim}{X}))|\mathcal{L}\} = 0, \quad \text{unless} \quad \underset{\sim}{X} = \gamma(\underset{\sim}{X}),$$

for all permutations γ in S' and all $\mathcal{L}$ in Ω'.

(ii) Let $R(h(\underset{\sim}{X})) = \sum_{\gamma\varepsilon S'} \varepsilon\{h(\underset{\sim}{X}) - h(\gamma(\underset{\sim}{X}))$, where

$\varepsilon(u) = 1$ if $u \geq 0$; $= 0$, if $u < 0$. Then, if $h(\cdot)$ is a B-Pitman function wrt S' and Ω', $R(h(\cdot))$ is called a B-Pitman statistic wrt S' and Ω'.

<u>Example 4.2.</u> Let $n = 3$, and $h(x_1, x_2, x_3) = x_1 + 2x_2 + 3x_3$. Then, $h(\cdot)$ is a B-Pitman function wrt S_3^* and $\Omega(SISI)$. Further, $R(h(\cdot))$ is a B-Pitman function wrt S_3^* and $\Omega(SISI)$.

One can prove immediately

<u>Theorem 4.3.</u> (Maximal Statistical Noise Theorem) (a) If $h(\cdot)$ is

a B-Pitman function wrt S_n^* and Ω(SISI), then $R(h(\cdot))$ is (a version of) M-S-N.

(b) T^* is NPDF wrt Ω(SISI) iff there exist a (measurable) function $W(\cdot)$ and B-Pitman function h^* such that $T^*(\underset{\sim}{X}) = R(h^*(\underset{\sim}{X}))$.

Corollary 4.1. (a) There are an infinite number of (mutually equivalent) versions of M-S-N.

(b) $N_1(\underset{\sim}{X})$, of Section 3, can be epxressed in terms of at least one B-Pitman statistic; and

(c) If $\mathcal{L} \in \Omega(\text{SISI})$, $P\{R(h(\underset{\sim}{X})) = k\} = [(n!)(2^n)]^{-1}$ for $1 \leq k \leq (n!)(2^n)$.

At this point one has the following formulations of basic statistical structure

M-S-S

(i) $S_1(\underset{\sim}{X}) = [|X|(1), \ldots, |X|(n)]$

(ii) $S_2^*(\underset{\sim}{X}) = \{\gamma(\underset{\sim}{X}): \gamma \in S_n^*\}$

(iii) $S_3(\underset{\sim}{X}) = G_n^*(\cdot)$ (of Def. 4.3 below)

M-S-N

(i) $N_1(\underset{\sim}{X}) = [\varepsilon(X_1), \ldots, \varepsilon(X_n); R(|X_1|), \ldots, R(|X_n|)]$

(ii) $N_h(\underset{\sim}{X}) = R(h(\underset{\sim}{X}))$, where $h(\cdot)$ is a B-Pitman function wrt S_n^* and Ω(SISI).

In order to make use of some standard nonparametric procedures, one introduces the entities below.

Definition 4.3. Let $F(\cdot)$ satisfy, $F(x) + F(-x) = 1$ for all x.

(a) $G_F(\cdot)$ is defined by $G_F(z) = 0$, for $z < 0$, and $=2F(z) - 1$ for $z > 0$.

(b) $G_n^*(\cdot)$ is defined by $G_n^*(z) = n^{-1} \sum_1^n (z - |X|(i))$.

One proves easily

Theorem 4.4. $S_3(\underset{\sim}{X}) = G_n^*(\cdot)$ is a M-S-S.

Further, it follows from Birnbaum and Rubin (1954), and Bell (1964a, b), that

Theorem 4.5. (PDF Theorem) Each statistic of the form

$$\phi[G_F(|X|(1)), \ldots, G_F(|X|(n))] \quad \text{or} \quad \psi^*(G_n^*(\cdot))$$

is PDF wrt $\Omega(\text{SISI})$.

This theorem will be used in constructing all procedures for one of the signal detection models of the sequel.

One final statistical tool will be introduced in this section. It is a slight generalization of the method of Durbin (1961) and Bell and Doksum (1965), and should be employed primarily to avoid certain distribution problems.

Let $\delta(\underset{\sim}{X}) = [S(\underset{\sim}{X}), N(\underset{\sim}{X})]$, where $\delta(\cdot)$, $S(\underset{\sim}{X})$, and $N(\underset{\sim}{X})$ are respectively, the BDT, M-S-S and M-S-N of a family Ω' of stochastic process laws. Let $\underset{\sim}{X}$ be data governed by a law $\mathcal{L}$ (usually unknown) in Ω', and let $\underset{\sim}{Y}$ be generated by a law $\mathcal{L}_0$ (known) in Ω' and be independent of $\underset{\sim}{X}$.

Definition 4.4. (a) $\underset{\sim}{Y}$ is called E-S-N (Extraneous statistical noise).

(b) $\underset{\sim}{Y}' = \delta^{-1}[S(\underset{\sim}{X}), N(\underset{\sim}{X})]$ is called R-S-N (randomized statistical noise).

Example 4.3. Let $\underset{\sim}{X} = (X_1, \ldots, X_{25})$ be the increments of data governed by a law $\mathcal{L}_F$ in Ω(SISI). Let $\underset{\sim}{Y} = (Y_1, \ldots, Y_{25})$ be the increments of a Wiener process (See Section 2) with $\sigma = 1/4$. Then $X_1, \ldots, X_{25}$ are i.i.d. $F(\cdot)$, unknown; and $Y_1, \ldots, Y_{25}$ are i.i.d. Φ. One simple version of the M-S-S is $S_1(\underset{\sim}{X}) = (|X|(1), \ldots, |X|(25)$, in which case one chooses $N_1(\underset{\sim}{X}) = [\varepsilon(X_1), \ldots, \varepsilon(X_{25});\ R(|X_1|), \ldots, R(|X_{25}|)]$, and, then $\delta(\underset{\sim}{X}) = [S_1(\underset{\sim}{X}), N_1(\underset{\sim}{X})]$. One then, forms $\delta^{-1}(|Y|(1), \ldots, |Y|(25); \varepsilon(X_1), \ldots, \varepsilon(X_{25});\ R(|X_1|), \ldots, R(|X_{25}|)) = (Y_1', \ldots, Y_{25}')$, where $Y_j' = [2\,\varepsilon\,(X_j) - 1]|Y_j|(R(|X_j|))$. This means each X is replaced by that Y, which has the same absolute-value rank and same sign. Then, it can be proved that $Y_1', \ldots, Y_{25}'$ are i.i.d. $N(0,1)$.

These ideas are formalized in the theorem below.

Theorem 4.6. (Randomized Noise Theorem). Let $\underset{\sim}{X} = (X_1, \ldots, X_n)$ and $\underset{\sim}{Y} = (Y_1, \ldots, Y_n)$ be independent, and be generated by laws $\mathcal{L}$ and $\mathcal{L}^*$, respectively in Ω'. Let $\delta(\cdot)$, $S(\underset{\sim}{X})$ and $N(\underset{\sim}{X})$, be respectively, the BDT, M-S-S and M-S-N for Ω'. If $\underset{\sim}{Y}' = (Y_1', \ldots, Y_n') =$

$\delta^{-1}[S(\underset{\sim}{Y}), N(\underset{\sim}{X})]$, then $Y' \overset{d}{=} Y \sim \mathcal{L}^* \varepsilon \Omega'$.

Thus, the distribution of the E-S-N, $\underset{\sim}{Y}$, has been imposed on the data while the M-S-N of the data is preserved.

Now, one is in a position to treat the pertinent inference problems.

5. Goodness-of-fit Detectors

The situation to be treated here is

$$\text{PN:}\quad \mathcal{L} = \mathcal{L}_0 \ \varepsilon\ \Omega(\text{SISI}) \quad \text{vs} \quad N + S \quad \mathcal{L} \neq \mathcal{L}_0.$$

One notes that the law $\mathcal{L}_0$ and the sampling spacing Δ, uniquely determine the common distribution F_0, of the independent symmetric increments $X_j = Z(j\Delta) - Z([j-1]\Delta)$, for $j = 1, 2, \ldots, n$, where $Z(0) = 0$.

Example 5.1. (a) In Example 2.1, $\mathcal{L}_0$ is such that $F_0 = N(0, \sigma^2\Delta)$. (b) In Example 2.3, if $\Delta = 3$, then F_0 is that distribution function with characteristic function $\phi(u) = (1 + t^2/\lambda)^{-3}$. In Example 2.2., F_0 is the distribution function with characteristic function $\phi(u) = \exp\{-\lambda\Delta[1 - \phi^*(u)]\}$, with $\phi^*(u) = (u\theta)^{-1} \sin u\theta$.

For a given constant sampling spacing Δ, one can rephrase the detection problem as

$$\text{PN}^*:\quad F = F_0 \ \varepsilon\ \Omega^* \quad \text{vs} \quad N + S:\ F \neq F_0$$

where $F_0(x) + F_0(-x) = 1$ for all x, i.e., F_0 exhibits symmetry wrt 0.

This being the case, one should base the decision rule on the M-S-S:

$S_1(\underset{\sim}{X}) = [|X|(1), \ldots, |X|(n)]$, and a PDF statistic using the cpf G_{F_0}, where $G_F(x) = 2F(x) - 1$, $x > 0$.

Hence (by Bell (1964b) and Birnbaum and Rubin (1954)), one should consider decisions based on statistics of the form $\psi[G_{F_0}(|X|(1)), \ldots, G_{F_0}(|X|(n))]$.

Decision Rule 5.1. Decide N + S iff $\sup_z G_n^*(z) - G_{F_0}(z) > d'$, where $G_n^*(z) = n^{-1} \sum_{j=1}^{n} \varepsilon\,(z - |X|(j))$, and $d' = d'(\alpha,n)$ chosen so as to achieve FAR,α.

Decision Rule 5.2. Decide N + S iff $\int [G_n^*(z) - G_{F_0}(z)]^2 dG_{F_0}(z) > w^* = w^*(\alpha,n)$.

Of course, several other goodness-of-fit statistics could have been used. [See, e.g. Bell (1964b), Model I].

If one knows that the signal has the effect of yielding a specified distribution, H for $\underset{\sim}{X}^* = (X_1^*, \ldots, X_k^*)$, where $X_j^* = |X|(j)$, then one would prefer a decision rule of the form below. [The problem here would be PN: $\mathcal{L} = \mathcal{L}_0$ vs N + S: $\mathcal{L} = \mathcal{L}_1$].

Decision Rule 5.3. Decide N + S iff $h(x_1^*, \ldots, x_k^*) > C^* \prod_1^k g_{F_0}(x_j^*)$, where $C^* = C(\alpha,k)$ and $g_{F_0}(\cdot)$ and $h(\cdot)$ are the appropriate densities.

These three decision rules will be illustrated with data in the appendix.

6. Class-Fit Detectors.

Here one has the detection problem PN: $\mathcal{L} \in \Omega(\mathrm{SISI})$ vs N + S: $\mathcal{L} \notin \Omega(\mathrm{SISI})$. Since this problem does not concern a specific value of the parameter (or law), one is led to NPDF statistics and M-S-N. One useful version is $N_1(\underset{\sim}{X}) = [\varepsilon(X_1), \ldots, \varepsilon(X_n); R(|X_1|), \ldots, R(|X_n|)]$.

Decision Rule 6.1. Decide N + S iff $T_n^*(\underset{\sim}{X}) = \sum_1^n \varepsilon(X_j) \geq c_1$ or $\leq c_2$, where $c_j = c_j(\alpha, n)$.

This, of course, is a Sign Detector or a Threshhold Detector. It makes no use of the rank vector.

In order to make use of all of the components of $S_1(\underset{\sim}{X})$, one introduces the Wilcoxon 1-sample statistic.

Decision Rule 6.2. Decide N + S iff $W_n^*(\underset{\sim}{X}) = \sum_1^n \varepsilon(X_j) R(|X_j|) \leq k_1$ or $\geq k_2$, where $k_j = k_j(\alpha, n)$.

The statistics employed in the two decision rules above involve statistics which are well-tabulated. However, these statistics both have discrete distributions, and it is necessary to turn to randomized procedures if one wishes to achieve certain particular PFA's,

One such procedure is based on E-S-N (extraneous statistical noise), i.e., $\underset{\sim}{Y} = (Y_1, \ldots, Y_n)$ independent of the data and generated by a known law $\mathcal{L}$ in $\Omega(\mathrm{SISI})$.

Let $Y_1, \ldots, Y_n$ be i.i.d. $N(0,1)$. This corresponds to data $\underset{\sim}{Z} = [Z(\Delta), Z(2\Delta), \ldots, Z(n\Delta)]$ from a WLP satisfying $\mu(t) \equiv 0$ and $\sigma^2\Delta = 1$. When the process $\{Z(t)\}$ is independent of the data so are the Y's.

From Section 4, one knows that the BDT, $\delta(\cdot)$, is such that $\delta(\underset{\sim}{X}) = [S_1(\underset{\sim}{X}), N_1(\underset{\sim}{X})]$, where $S_1(\underset{\sim}{X}) = [|X|(1), \ldots, |X|(n)]$ and $N_1(\underset{\sim}{X}) = [\varepsilon(X_1), \ldots, \varepsilon(X_n); R(|X_1|), \ldots, R(|X_n|)]$. In order to apply the Randomized Noise Theorem (Theorem 4.6), one forms $\underset{\sim}{Y}' = (Y_1', \ldots, Y_n') = \delta^{-1}[S_1(\underset{\sim}{Y}), N_1(\underset{\sim}{X})]$. In this case $Y_j' = [2\,\varepsilon(X_j)-1]|Y_j|(R(|X_j|))$. This means $N_1(\underset{\sim}{X}) = N_1(\underset{\sim}{Y}')$. By Theorem 4.6, $Y_1', \ldots, Y_n'$ are i.i.d. $N(0,1)$ under PN: $\mathcal{L} \in \Omega(\text{SISI})$. One useful decision rule based on the Kolmogorov-Smirnov statistic is then.

Decision Rule 6.3. Decide N + S iff $\sup_z \left|\frac{1}{n} \sum_1^n (z - Y_j') - \Phi(z)\right| > d'$, where $d' = d(\alpha, n)$.

Again one notes that Decision Rule 6.3. could have employed any goodness-of-fit statistic. (See, e.g., Bell (1964a); Model I Detectors)).

A different type of detection procedure is developed if one employs a different version of the M-S-N.

Let S_n^* (See Definition 4.1) be the set of sign-time permutations of coordinates of $\underset{\sim}{X} = (X_1, \ldots, X_n)$, and let $h(\underset{\sim}{X}) = \sum_1^n j\, X_j$. Then one has the following result.

Theorem 6.1. (a) $h(\cdot)$, above, is a B-Pitman function wrt $\Omega(\text{SISI})$ and S_n^*.

(b) $R(h(\underset{\sim}{X}))$, as in Definition 4.1, is M-S-N.

(c) Under PN, $R(h(\underset{\sim}{X})) \sim D - U\{1, 2, \ldots, k^*\}$, i.e., has a discrete uniform distribution over the integers $\{1, 2, \ldots, k^*\}$, where $k^* = (n!)(2^n)$.

Decision Rule 6.4. Decide N + S iff $R(h(\underset{\sim}{X})) \leq b_1$ or $\geq b_2$. where $b_1 = (n!)(2^n)(\frac{\alpha}{2})$ and $b_2 = 1 + (1 - \frac{\alpha}{2})(n!)(2^n)$.

Of course, one could construct an analagous decision rule for each B-Pitman function.

These decision procedures are illustrated in the appendix.

The final detection problem of this paper concerns the equality of two stochastic process laws.

7. Two-Sample Detectors.

The detection problem considered here is

$$\text{PN: } \mathcal{L}_1 = \mathcal{L}_2 \quad \text{vs} \quad \text{N + S: } \mathcal{L}_1 \neq \mathcal{L}_2$$

where $\mathcal{L}_j \in \Omega(\text{SISI})$.

Consider the situation where $\underset{\sim}{X} = (X_1, \ldots, X_m)$ is generated by $\mathcal{L}_1$ and $\underset{\sim}{Y} = (Y_1, \ldots, Y_n)$ is generated by $\mathcal{L}_2$ and is independent of $\underset{\sim}{X}$.

Under PN, $\underset{\sim}{Z} = (\underset{\sim}{X},\underset{\sim}{Y}) = (X_1, \ldots, X_m; Y_1, \ldots, Y_n) = (Z_1, \ldots, Z_N)$ is distributed as a random sample from an (unknown) $F(\cdot)$, symmetric wrt 0.

Since no specific value of a parameter (or law) is involved, one needs an NPDF statistic. Such a statistic must be based on some version of the M-S-N, e.g., $S_1(\underset{\sim}{Z}) = [\varepsilon(Z_1), \ldots, (Z_N); R(Z_1), \ldots, R(Z_N)]$.

Decision Rule 7.1. Decide N + S iff $W = \sum_1^n R(|Z_j|) \geq b_1$ or $\leq b_2$, where $b_j = b_j(\alpha,m,n)$.

Decision Rule 7.2. Decide N + S iff

$$\sup_{z} \left| \frac{1}{m} \sum_{1}^{m} (z - |X|(j)) - \frac{1}{n} \sum_{1}^{m} (z - |Y|(j)) \right| > d' = d(\alpha,m,n)$$

Decision Rule 7.3. Decide N + S iff

$$\frac{1}{m} \sum_{1}^{m} V(R|X_j|) - \frac{1}{n} \sum_{1}^{n} V(R(|Y_j|)) > a_1 \quad \text{or} \quad < a_2,$$

where $V_1, \ldots, V_n$ are i.i.d. Φ and independent of the data.

This last decision rule is based on the Theorem 4.6. (See also, Bell and Doksum (1965)). The first and second rules above correspond to the Wilcoxon Rank-Sum statistic, and two-sample Kolmogorov-Smirnov statistic, respectively.

The techniques here, although concerned with Ω(SISI) are closely related to many of the usual two-sample non-parametric detectors. (See, e.g., Bell (1964a; Model II Detectors)).

Numerical examples of the decision rules here are given in the appendix. We conclude this section with two more detectors.

(A) A Modified Two-Sample Detector

The simplest rank detector (statistic) for the pure noise situation of univariate symmetry about zero is: $T = \sum_{i=1}^{N} \varepsilon(Z_i)$ where $\varepsilon(y) = 1$ if $y \geq 0$ and zero otherwise. Under the null hypothesis of PN, T has a Binomial distribution with parameters N and $\frac{1}{2}$. However, it is clear that for any asymmetric distribution with $F(0) = \frac{1}{2}$, the power of any test based solely on T is equal to its PFA.

Let $T = k$, and define $\{Z_1', \ldots, Z_{N-k}'\} = \{|Z_i|: Z_i < 0\}$ and

$\{Z_1^*, \ldots, Z_k^*\} = \{|Z_i|: Z_i > 0\}$. There are now two new samples on which one can try any two-sample detection procedure such as the Kolmogorov-Smirnov detector, Cramer-von Mises detector, Mann-Whitney-Wilcoxon detector, Fisher-Yates-Terry detector etc. Each of these tests is a most powerful invariant detector against specific $N + S$ alternatives, see Lehmann (1959).

The power of such tests is equal to PFA for $S + N$ situations with the Z distribution symmetric about zero. There is a slight difficulty in applying detectors based on comparisons of $\{-Z_j'\}$ and $\{Z_j^*\}$ when m or n is small. For this latter reason, one can introduce the modification below. (The development below is a slight extension of some ideas in: Dippo, C. (1970). Distribution-free tests of univariate symmetry about zero, Class Project, University of Michigan.)

Let $T(m,n)$ be an arbitrary two-sample detector with PFA α and critical region $C(\alpha;m,n)$. Then let $\psi(\underset{\sim}{Z}) = 1$ if $T < k_1$ or $> k_2$; $= 1$ if $k_1 \leq T \leq k_2$ and $T(m,n) \in C(\alpha;m,n)$; and zero otherwise. The power of detector ψ at $F_1 = P\{\text{Reject PN} | S + N \text{ is true}\}$ is

$$\int \psi dF_1 = P\{T < k_1 | F_1\} + P\{T > k_2 | F_1\} +$$

$$+ \sum_{j=k_1}^{k_2} P\{T = j\} \, P\{T(j,N-j) \in C(\alpha;j,N-j)\}$$

$$= \sum_{j=1}^{k_1-1} {}^N C_j \, p_1^j \, q_1^{N-j} + \sum_{j=k_2+1}^{N} {}^N C_j \, p_1^j \, q_1^{N-j} +$$

$$+ \sum_{j=k_1}^{k_2} {}^N C_r \, P\{T(j,N-j) \in C(\alpha;j,N-j)\} \, p_1^j \, q_1^{N-j} ,$$

where $p_1 = P\{Z < 0|F_1\}$ and $q_1 = 1 - p_1$.

This procedure eliminates the difficulties when m or n is small, but the power is still α = PFA when one has univariate symmetry about zero - a desirable property.

(B) A Chi-Squared Detector

Under the PN situation, divide the domain of Z_j's into sections symmetric about zero by the set of numbers $\{a_j, j = 0, \pm 1, \pm 2, \ldots, \pm k\}$. The probability of falling in the interval (a_{j-1}, a_j), $j \geq 1$ under the PN case is the same as falling in the interval $(-a_j, -a_{j-1})$. The a_j's may be chosen arbitrarily, be based on some prior knowledge, or, better yet, be based on the ordered statistics.

Let $Z_{(1)}, \ldots, Z_{(N)}$ be the order statistics of the $\{|Z_j|\}$. If the desired number of intervals is k, let $r_j = jN/k$ and $a_j = Z_{(r_j)}$, $j = 1, 2, \ldots, k$. Then

$$N_{1j} = \sum_{i=1}^{N} \varepsilon(Z_i)[\varepsilon(a_j - Z_i) - \varepsilon(a_{j-1} - Z_i)]$$

$$N_{2j} = \sum_{i=1}^{N} \varepsilon(-Z_i)[\varepsilon(a_j - Z_i) - \varepsilon(a_{j-1} - Z_i)].$$

Let $N_{1\cdot} = \sum_j N_{1j}$ and $N_{2\cdot} = \sum_j N_{2j}$. Therefore, under the PN situation, the detector

$$\sum_i \sum_j \left[N_{ij} - \frac{N_{1\cdot}N_{2\cdot}}{N}\right]^2 \left[\frac{N}{N_{1\cdot}N_{2\cdot}}\right]$$

is DF, and has an asymptotic null distribution of a Chi-squared statistic with $(k - 1)$ degrees of freedom.

8. Point Estimation and Confidence Bounds for $F_{\mathcal{L}}$ with $\mathcal{L}$ in $\Omega(\mathrm{SISI})$.

(A) Point Estimation

Recall the definitions of SISI and symmetric (about zero) probability laws, that is

$$\Omega(\mathrm{SYM}) = \{F(\cdot):\ F \text{ is continuous and } F(x) + F(-x) = 1 \text{ for all } x\}$$

$$\Omega(\mathrm{SISI}) = \{\mathcal{L}:\ Y_1, \ldots, Y_m \text{ are i.i.d. } F \in \Omega(\mathrm{SYM}) \text{ for all } \Delta > 0, \text{ where } Y_j = Z(j\Delta) - Z((j-1)\Delta),\ Z(0) = 0\}.$$

Example 8.1. Let $X_1, X_2, \ldots, X_m, \ldots$ be i.i.d. Cauchy with parameter $(0,\theta)$, $C(0,\theta)$, that is $f(x) = \dfrac{\theta}{\pi(\theta^2 + x^2)}$, $-\infty < x < \infty$; and let $Z_n = \sum_{j=1}^{n} X_j$. Choose $\Delta = 3$ and set $Y_1 = Z_3$, $Y_2 = Z_6 - Z_3$, $Y_3 = Z_9 - Z_6, \ldots$ Then, for $\{Z_m:\ m \geq 1\}$, (a) $\mathcal{L} \in \Omega(\mathrm{SISI})$; and (b) $Y_1, \ldots, Y_m$ are i.i.d. $C(0, 3\theta)$ which belongs to $\Omega(\mathrm{SYM})$.

To estimate $F(x^*)$ for $x^* > 0$, clearly a natural estimator is

$$\hat{F}(x^*) = F_m(x^*) = m^{-1} \sum_{j=1}^{m} \varepsilon\,(x^* - X_j).$$

If $x^* < 0$, use estimate $1 - F(-x^*)$. Obviously,

$$E\{\hat{F}(x^*)\} = F(x^*),\quad \mathrm{Var}\{\hat{F}(x^*)\} = \frac{F(x^*)[1-F(x^*)]}{m} = \frac{F(x^*)F(-x^*)}{m},$$

and $m\hat{F}(x^*)$ is a Binomial variable with parameters $(m, F(x^*))$.

An optimal estimate for $F(x^*)$ can be obtained by using only the M-S-S, $\underset{\sim}{S}(Z) = \{|X|(1), \ldots, |X|(m)\}$ or another, easier to handle, version of it $G_m^*(\cdot)$, where $G_m^*(z) = m^{-1} \sum_{j=1}^{m} \varepsilon\,(z - |X|(j))$.

<u>Lemma 8.1</u>. (a) $P\{|X_1| \leq z\} = 2F(z) - 1$, $z \geq 0$ and zero otherwise,

(b) $E\{G_m^*(x^*)\} = 2F(x^*) - 1$;

(c) $\mathrm{Var}\{G_m^*(x^*)\} = \dfrac{2[1 - F(x^*)][2F(x^*) - 1]}{m}$.

On the basis of the definition of $G_m^*(\cdot)$ and the above lemma, define another estimator of $F(x^*)$ as

$$\tilde{F}_m(x^*) = \frac{G_m^*(x^*) + 1}{2}$$

<u>Theorem 8.1</u>. (a) $E\{\tilde{F}_m(x^*)\} = F(x^*)$;

(b) $\mathrm{Var}\{\tilde{F}_m(x^*)\} = \dfrac{[1 - F(x^*)][2F(x^*) - 1]}{2m}$

(c) $\tilde{F}_m(x^*)$ is the UMVU estimator for $F(x^*)$.

(d) $0 \leq \dfrac{\mathrm{Var}\{\tilde{F}(x^*)\}}{\mathrm{Var}\ \hat{F}(x^*)} = \dfrac{2F(x^*) - 1}{2F(x^*)} \leq \dfrac{1}{2}$

Thus, one can do better by using the estimator $\tilde{F}(x^*)$ which uses the symmetry property of the underlying process.

(B) <u>Confidence Bounds</u>

In this section the aim is to establish that when one knows the underlying distribution function is symmetric about zero, then one can achieve a confidence region having 50% less interval length as compared to the usual Kolmogorov-Smirnov-type confidence regions. To put it in another way, it means that with the same confidence level $(1 - \alpha)100\%$ of the former confidence interval has half the width than that of the latter. First we state some relevant results which can easily be proved.

<u>Theorem 8.2.</u> (a) For $Y_1, \ldots, Y_m$ i.i.d. with $F \in \Omega(\text{SYM})$, the M-S-S is $S(\underset{\sim}{Y}) = [|Y|(1), \ldots, |Y|(m)]$, that is the ordered absolute values.

(b) $|Y_j| \sim G_F$, where

$$G_F(z) = \begin{cases} 2F(z) - 1, & \text{for } z \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

(c) $$F(z) = \begin{cases} \frac{1}{2}[1 + G_F(z)], & z \geq 0 \\ \frac{1}{2}[1 - G_F(-z)], & z < 0 \end{cases}$$

<u>Definition 8.1.</u> $\Omega_2(R_1^+) = \{G\colon\ G \text{ is continuous},\ G(0) = 0\}$.

<u>Theorem 8.3.</u> F belongs to $\Omega(\text{SYM})$, if and only if, G_F is a member of the class $\Omega_2(R_1^+)$.

<u>Definition 8.2.</u> Let $d(\alpha,m)$ be such that

$$\alpha = P\{\sup_z |H_m(z) - H(z)| > d(\alpha,m)\},$$

where $V_1, \ldots, V_m$ are i.i.d. with distribution $H(\cdot)$ which is continuous; and $H_m(z) = m^{-1} \sum_{j=1}^{m} \varepsilon\,(z - V_{(j)})$.

Let $Y_1, \ldots, Y_m$ be i.i.d. with $F \in \Omega(\mathrm{SYM})$. Then $|Y_1|, \ldots, |Y_m|$ are i.i.d. with $G_F \in \Omega_2(R_1^+)$.

Lemma 8.2. The following statement holds:

$$P\{G_m^*(z) - d(\alpha,m) < G_F(z) < G_m^*(z) + d(\alpha,m) \text{ for all } z\} = 1 - \alpha,$$

where $G_m^*(\cdot)$ is the empirical distribution of the $|Y|$'s.

Lemma 8.3. (a) A consequence of the above result and the previous development is that:

$$P\{\frac{1 + G_m^*(z)}{2} - \frac{1}{2}\,d(\alpha,m) < F(z) < \frac{1 + G_m^*(z)}{2} + \frac{1}{2}\,d(\alpha,m) \text{ for all } z\}$$

$$= 1 - \alpha = P(A^+), \text{ say.}$$

(b) $$P(A^+) = 1 - \alpha = P\,\{\frac{1 - G_m^*(y)}{2} - \frac{1}{2}\,d(\alpha,m) < F(y) < \frac{1 - G_m^*(y)}{2} + \frac{1}{2}\,d(\alpha,m)$$

$$\text{for all } y < 0\}$$

$$= P(A^-), \text{ say.}$$

(c) $P\{A^+\} = P\{A^-\} = P\{A^+ \cap A^-\} = 1 - \alpha.$

Definition 8.3. Let $L_m^*(z)$ and $L_m(z)$ be defined as follows

$$L_m^*(z) = \begin{cases} \frac{1}{2}\,[1 + G_m^*(z)] + \frac{1}{2}\,d(\alpha,m), & z \geq 0 \\ \frac{1}{2}\,[1 - G_m^*(-z)] + \frac{1}{2}\,d(\alpha,m), & z < 0 \end{cases}$$

$$L_m(z) = \begin{cases} \frac{1}{2}\,[1 + G_m^*(z)] - \frac{1}{2}\, d(\alpha,m), & z \geq 0 \\ \frac{1}{2}\,[1 - G_m^*(-z)] - \frac{1}{2}\, d(\alpha,m), & z < 0 \end{cases}$$

Theorem 8.4. Let $Y_1, \ldots, Y_m$ be i.i.d. with $F(\cdot)$ in $\Omega(\mathrm{SYM})$ and $G_m^*(z) = m^{-1} \sum_{j=1}^{m} \varepsilon\,(z - |Y|(j))$. Then

(a) $P\{L_m(z) < F(z) < L_m^*(z),$ for all $z\} = 1 - \alpha.$

(b) $L_m^*(z) - L_m(z) = d(\alpha,m)$ for all z.

The above result combined with the statement below helps achieve our aim set in the beginning of this section.

Theorem 8.5. Let $V_1, \ldots, V_m$ be i.i.d. with distribution $H(\cdot)$ which is continuous. Then

(a) $P\{H_m(z) - d(\alpha,m) < H(z) < H_m(z) + d(\alpha,m)$ for all $z\} = 1 - \alpha$

(b) $[H_m(z) + d(\alpha,m)] - [H_m(z) - d(\alpha,m)] = 2d(\alpha,m)$ for all z.

These results can be presented in a slightly different form when one constructs the following empiric-type distribution - the empirical distribution of data:

$$-|Y|(m),\ -|Y|(m-1),\ \ldots,\ -|Y|(1);\ |Y|(1),\ \ldots,\ |Y|(m)$$

Definition 8.4. Let $G_m^{**}(z)$ be as below:

$$G_m^{**}(z) = (2m)^{-1}\,[\sum_{j=1}^{m} \varepsilon\,(z - |Y|(j)) + \sum_{j=1}^{m} \varepsilon\,(z + |Y|(j))].$$

Lemma 8.4. With probability one,

(a) $G_m^{**}(z) = \frac{1}{2}\{1 + [G_m^*(z) - G_m^*(-z)]\}$, for all z; and

(b) $$G_m^{**}(z) = \begin{cases} \frac{1}{2}[1 + G_m^*(z)], & z \geq 0 \\ \frac{1}{2}[1 - G_m^*(-z)], & z < 0 \end{cases}$$

Theorem 8.6. Let $Y_1, \ldots, Y_m$ be i.i.d. with $F(\cdot)$ in $\Omega(SYM)$. Then,

$$P\{G_m^{**}(z) - \frac{1}{2} d(\alpha,m) < F(z) < G_m^{**}(z) + \frac{1}{2} d(\alpha,m) \quad \text{for all } z\}$$

$$= 1 - \alpha.$$

The results of this section "fit in" well with the idea that in constructing PDF (parametric distribution-free) statistics for $\Omega(SISI)$ one should use only the M-S-S, $S(\underset{\sim}{Z}) = [|Y|(1), \ldots, |Y|(m)]$. To use any "additional aspects of the data" should be inefficient.

REFERENCES

1. Ahmad, R. (1974). On the structure of symmetric sample testing: a distribution-free approach. Ann. Inst. Statist. Math., 26, 233-245. (1980).
2. Basawa, I. V. and Prakasa, Rao, B. L. S. (1980). Statistical Inference for Stochastic Processes. Academic Press, New York.
3. Bell, C. B. (1964a). Automatic Distribution-Free Statistical Signal Detection. Research Report, US Navy, Electronics Laboratory, San Diego.
4. Bell, C. B. (1964b). Some basic theorems for distribution-free statistics. Ann. Math. Statist., 35, 150-156.
5. Bell, C. B. and Doksum, K. J. (1965). Some new distribution-free statistics. Ann. Math. Statist., 36, 203-214.
6. Bell, C. B. and Haller, H. S. (1969). Bivariate symmetry tests: parametric and nonparametric. Ann. Math. Statist., 40, 259-269.
7. Bell, C. B. and Smith, P. J. (1969). Some nonparametric tests for the multivatiate goodness-of-fit, multisample, independence and symmetry problems. In "Multivariate Analysis", Academic Press, New York.
8. Bell, C. B., Woodroofe, M. F. and Avadhani, T. V. (1970). Some nonparametric tests for stochastic processes. In "Nonparametric Techniques in Statistical Inference", Cambridge University Press.

9. Bell, C. B. and Smith, P. J. (1972). Symmetric functions in statistics. In "Proc. Symp. in Honour of Professor P. S. Dwyer," University of Windsor, Canada.

10. Bell, C. B., Ramirez, F. and Smith, E. P. (1980). Wiener-Levy models spherically exchangeable time series, and simultaneous inference in growth curve analysis. In "Advanced Asymptotic Testing and Estimation: A Symposium in Honour of Wassily Hoeffding", Academic Press.

11. Birnbaum, Z. W. and Rubin, H. (1954). On distribution-free statistics. Ann. Math. Statist., 25, 593-598.

12. Conover, W. J. (1971). Practical Nonparametric Statistics. Wiley.

13. Doob, J. L. (1953). Stochastic Processes. Wiley, New York.

14. Dubrin, J. (1961). Some methods of constructing exact tests. Biometrika, 48, 41-55.

15. Feller, W. (1966). An Introduction to Probability Theory and its Applications, Vol. 2. Wiley, New York.

16. Hollander, M. and Wolfe, D. A. (1973). Nonparametric Statistical Methods. Wiley.

17. Kendall, M. G. (1976). Time-Series. Charles Griffin & Co. Ltd, London.

18. Lehmann, E. L. (1959). Testing Statistical Hypotheses. Wiley, New York.

APPENDIX FOR Ω(SISI) FAMILY

ILLUSTRATIVE EXAMPLES AND GRAPHS +

Example 1. To illustrate the Goodness-of-fit detection procedure we considered the data given by Dewey (1963). The graph of the data is shown in Graph 1(a). Figure (b) shows the power spectrum of smoothed U.S.A. immigration data of Dewey. By taking the period 1830 to 1950 we looked at residuals of logarithms of immagration data, after removal of trend by a simple seventeen-point moving average - values increased by 2, see Kendall (1976, p. 103). Let X_j denote the jth year residual value for the period considered, then set $W_1 = X_1 + \ldots + X_{10}$, $W_2 = X_{11} + \ldots + X_{20}$; ...; $W_{12} = X_{111} + \ldots + X_{120}$; here $\Delta = 10$ and $n = 1, 2, \ldots, 12$. This process gives a probability law which belongs to Ω(SISI) but is not symmetric about zero. The aim is to detect: - PN: $\mathcal{L} \in \Omega(\text{SISI})$ (with symmetry around 18) against N + S: $\mathcal{L} \equiv U_{(10,50)}$, U indicates uniform distribution. The ordered W_j's with computations and decision rule are presented below.

+This Appendix was prepared with the aid of S-M Lee and A. Mason of San Diego State University.

*Dewey, E. R. (1963). The 18·2-year cycle in immigration, U.S.A., 1820-1962. Foundation for the Study of cycles, Inc., Pittsburgh, PA.

Table 1: U.S.A. Immigration Data 1830-1950.

Ordered Data	Computations	
$W_{(j)}$	$\frac{j}{12} - \frac{W_{(j)}}{40}$	$\frac{W_{(j)}}{40} - \frac{j-1}{12}$
18.5484	-0.3804	0.4637
19.1244	-0.3114	0.3948
19.1562	-0.2289	0.3122
19.1853	-0.1463	0.2296
19.3680	-0.0675	0.1509
19.5932	0.0102	0.0732
19.8373	0.0874	-0.0041
20.4530	0.1553	-0.0720
20.7822	0.2304	-0.1471
20.8606	0.3118	-0.2285
21.0688	0.3899	-0.3066
22.0959	0.4476	-0.3643

The detection procedure used is the Kolmogorov-Smirnov statistic

$$D_{12} = \max_j \{\max [\frac{j}{12} - \frac{W_{(j)}}{40}, \frac{W_{(j)}}{40} - \frac{j-1}{12}]\}$$

$$= 0.4637$$

With PFA $\alpha = .01$, the critical value is $d_{12;005} = 0.449$, see Conovor (1971, p.397).

Decision Rule: Decide N + S since $D_{12} > d_{12;005}$.

Example 2. (Random Walk). To explain the detection procedures for the Class-Fit problem, 30 observations were generated from a Cauchy distribution with location and scale parameters, respectively, zero and 2. Thus, $Y_1, Y_2, \ldots, Y_{30}$ are i.i.d. r. vs. from $C(0,2)$. Let $Z(j) = \sum_{k=1}^{j} Y_k$, then $\{Z(j)\}$ is a stochastic process with law in the family $\Omega(SISI)$. Choose $\Delta = 3$, and set $X_1 = Z(3)$, $X_2 = Z(6) - Z(3)$, $\ldots$, $X_{10} = Z(30) - Z(27)$. Clearly, $\{Z(j)\}$ forms a random walk and $\{X_j\}$ is a sample from it. For the plot of the data see Graph 2.

Table 2. Random Walk Data

Data		Computations		
		X_i	$\varepsilon(X_i)$	$R(\lvert X_i\rvert)$
Y_1 = 0.5143	Y_{16} = -97.8398	15.7933	1	6
Y_2 = 13.2748	-105.2880	-103.4208	0	10
15.7933	-39.1513	-1.7487	0	1
14.3173	-8.3701	-1.8538	0	2
14.0243	-6.9749	-7.4049	0	4
-87.6275	2.6035	59.4836	1	9
-89.3277	0.0598	41.7548	1	8
-90.7272	-4.5361	-2.7977	0	3
-89.3762	-0.1942	-30.4395	0	7
-90.1488	2.9335	11.6435	1	5
-94.6573	3.1908			
-91.2300	-30.6337			
-90.6965	-31.9554			
-96.9435	-29.7549			
Y_{15} = -98.5349	Y_{30} = -18.9902			

To detect PN: $\mathcal{L} \in \Omega(SISI)$ versus $\mathcal{L} \notin \Omega(SISI)$, the procedure is as given below.

The Sign Detector. One proceeds as follows:

$$T^*_{10}(\underset{\sim}{X}) = \sum_{j=1}^{10} \varepsilon\,(X_j) = 4$$

$$P\{T^*_{10} \leq 1\,;\, n = 10\} = 1 - P\{T^*_{10} \geq 8;\ n = 10\} = 0.0107$$

Therefore, $\alpha = 0.0214$; and one decides $N + S$, if and only if, $T^*_{10}(\underset{\sim}{X}) \geq 8$ or ≤ 1. The conclusion is decide PN, with PFA $= \alpha = .01$.

The Wilcoxon-Signed-Rank Detector. In this situation one computes $W^*_{10}(\underset{\sim}{X}) = \sum_{j=1}^{10} \varepsilon\,(X_j)\ R(|X_j|) = 28$. We decide $N + S$, if and only if $W^*_{10}(\underset{\sim}{X}) \geq k_1$ or $\leq k_2$, where

$$k_1 = t(\alpha/2,\ n) = t(.005,10) = 52$$

$$k_2 = \frac{n(n+1)}{2} - t(\alpha/2,\ n) = 3$$

("t" denotes "tabulated value" obtained from Hollander and Wolfe (1973)). Since W^*_{10} is outside the critical region, one decides PN.

Example 3. (Compound Poisson Process). Let $Z(t) = \sum_{j=1}^{N(t)} T_j$, where $\{T_j\}$ are i.i.d. r. vs. from double exponential with parameter 4; $\{N(t)\}$ is homogeneous Poisson process with parameter 0.5; and $\{T_j\}$ and $\{N(t)\}$ are independent. Table 3 below gives the generated data from this process. The plot for jumps is shown is graph 3.

We sample every three units, that is the basic data is $\{Z(3), Z(6), \ldots, Z(30)\}$. Let $X_1 = Z(3)$, $X_2 = Z(6) - Z(3)$, ..., $X_{10} = Z(3) - Z(27)$. The interest here is to detect PN: $\mathcal{L}_X \in \Omega(\text{SISI})$ against N + S: $\mathcal{L} \notin \Omega(\text{SISI})$. The appropriate detection statistic is the Wilcoxon one-sample statistic. The essential computations are as follows:

Compound Poisson Data

j	$Z(3j)$	$X_j=Z(3j)-Z(3(j-1))$	$R(\|X_j\|)$
1	0.4576	0.4576	5
2	1.3184	0.8608	8
3	2.1871	0.8687	9
4	1.7750	-0.4121	4
5	2.7044	0.9294	10
6	3.0626	0.3582	3
7	3.1917	0.1291	1
8	2.7303	-0.4614	6
9	2.4915	-0.2388	2
10	3.0675	0.5760	7

<u>Decision Rule</u>. Since $W^*_{12} = \sum_{j=1}^{12} \varepsilon(X_j)R(|X_j|) = 43$ does not fall in the critical region $\{W^*_{12} \geq 52$ or $W^*_{12} \leq 3\}$ with PFA $\alpha = .01$, we decide PN.

Table 3. Compound Poisson Process Data

Interarrival Times	"Jumps"	Cumulative Jumps
1.0895	0.1250	0.1250
0.1005	0.5383	0.6633
0.4816	-0.2057	0.4576
2.4213	-0.0637	0.3939
1.5793	-0.0065	0.3874
10.2320	0.9310	1.3184
2.5747	0.3350	1.6534
2.3683	0.0680	117214
0.7450	0.4657	2.1871
1.9194	-0.1040	2.0831
4.0348	-0.2088	1.8743
0.3678	-0.0993	1.7750
1.0791	0.0779	1.8529
4.6253	0.9813	2.8342
2.5675	-0.1298	2.7044
6.9561	-0.0153	2.6891
4.9539	0.1758	2.8649
0.0201	0.1977	3.0626
0.0424	0.0324	3.0950
0.7306	-0.1454	2.9496
0.1366	0.2421	3.1917
3.1023	-0.2848	2.9069
4.0651	-0.2780	2.6289
0.2947	-0.1014	2.7303
0.3993	-0.0002	2.7301
1.2287	-0.3191	2.4110
7.9266	0.0805	2.4915
2.3167	0.3073	2.7988
0.5358	0.0422	2.8410
0.1195	0.2265	3.0675

Example 4 (Wiener-Levy Process). Let $W(t) = \mu(t) + \sigma W^*(t)$, where $\{W^*(t), t \geq 0\}$ is a Wiener-Levy process (WLP) satisfying (i) $EW^*(t) \equiv 0$; (ii) $\mathrm{Cov}\{W^*(s), W^*(t)\} = \min(s,t)$; (iii) it is Gaussian. Let $Z_j = W(j\Delta)$, that is one samples at times $\Delta, 2\Delta, \ldots,$ and $\underset{\sim}{X} = (X_1, \ldots, X_m)$ where $X_j = Z_j - Z_{j-1}$ and $Z_0 = 0$. If $\mu(t) = 0$, the family of such processes is denoted by $\Omega(\mathrm{WLP}\phi)$; and if $\mu(t) = t$, then it is denoted by $\Omega(\mathrm{WLPL})$.

In this example we illustrate the 2-sample detection procedure for PN: $\mathcal{L}_1 = \mathcal{L}_2 \in \Omega(\mathrm{SISI})$ against N + S: $\mathcal{L}_1 \neq \mathcal{L}_2$, $\mathcal{L}_i \in \Omega(\mathrm{SISI})$ for $i = 1, 2$.

For $\mathrm{WLP}\phi$, we chose $(\mu,\sigma,\Delta) = (0, 5, 0.1)$ and generated the first set of data $(X_1, \ldots, X_m)$ $(m = 15)$ as follows:

(1) Generate 30 $N(0,1)$ observations,

(i) Generate $\theta_1, \ldots, \theta_{15}$ $U(0, 2\pi)$ observations,

(ii) Generate $R_1^2, \ldots, R_{15}^2$ Exp. (1/2) observations.

(2) Let $X^*_{2m-1} = R_m \cos\theta_m$, $X^*_m = R_m \sin\theta_m$.

(3) $X'_j = 5\sqrt{\Delta}\, X^*_j$ gives $X'_1, \ldots, X'_{30}$ i.i.d. r. vs. from $N(0, 25\Delta)$.

(4) $X'_1 = Z(\Delta)$, $X'_2 = Z(2\Delta) - Z(\Delta)$, $\ldots$, $X'_{30} = Z(30\Delta) - Z(29\Delta)$ gives $Z(j\Delta) = \sum_{i=1}^{j} X'_i$, $j = 1, 2, \ldots, 30$. This provides first-sample $(X_1, X_2, \ldots, X_{15})$. By choosing $(\mu(t),\sigma,\Delta) = (3t, 5, \Delta = 0.5)$, we generate similarly the second sample of observations (from WLPL) which is denoted by $(Y_1, Y_2, \ldots, Y_{15})$. Thus the BDT is $Z = (X_1, \ldots, X_{15};$

$Y_1, \ldots, Y_{15}$). The probability laws for X and Y, respectively, are denoted by $\mathcal{L}_1$ and $\mathcal{L}_2$. The plots of $\{X_j'\}$ and $\{Z(j\Delta)\}$ for both processes are shown in graphs 4 to 7. The Wilcoxon Rank-Sum Detector is given below; and the Kolmogorov-Smirnov Detector is presented in Table 4.

(A) Wilcoxon Rank-Sum Detector. The data is as given in the first two columns of Table 4. By using $\{X_j\}$ we compute:

$$\{R(|X_j|),\ j = 1,2,\ \ldots,\ 15\} = \{1,2,3,4,6,8,10,11,12,14,15,16,17,\ 18,\ 21\}$$

Detector: $$W = \sum_{j=1}^{15} R(|X_j|) = 158.$$

Using a large sample approximation for $m = n = 15$, one gets

$$W^* = \frac{W - [n(m + n + 1)/2]}{[mn(m + n + 1)/12]} = -3.09; \quad W^* \sim N(0,1) \text{ asymptotically.}$$

Decision Rule: Since $W^* < Z_{\alpha/2} = -1.96$, we conclude $N + S$.

The conclusion is still the same if one chooses PFA $\alpha = .01$.

(B) Two-Sample Kolmogorov-Smirnov Detector for Wiener Levy Data

Table 4. Computations

Ordered Combined Data: $\lvert Z\rvert(j)$		$\left\lvert (15)^{-1}\sum_{j=1}^{15}\varepsilon(z-\lvert X\rvert(j))-(15)^{-1}\sum\varepsilon(z-\lvert Y\rvert(j))\right\rvert$
$\lvert X\rvert(j)$	$\lvert Y\rvert(j)$	
0.0680	-	1/15
0.1644	-	2/15
0.2135	-	3/15
0.4502	-	4/15
-	0.499	3/15
0.6046	-	4/15
-	0.7744	3/15
0.9073	-	4/15
-	0.9714	3/15
1.1185	-	4/15
1.6135	-	5/15
1.6879	-	6/15
-	1.6923	5/15
1.9042	-	6/15
2.0984	-	7/15
2.2764	-	8/15
2.2999	-	9/15
2.4765	-	10/15
-	2.6325	9/15
-	2.8479	8/15
3.285	-	9/15
-	3.4773	8/15
-	4.0665	7/15
-	4.3455	6/15
-	4.3518	5/15
-	6.6080	4/15
-	7.2577	3/15
-	8.0898	2/15
-	8.1427	1/15
-	8.5778	0

$m = n = 15$

$D_{m,n} = \sup_{z} \lvert F_m(z)-G_n(z)\rvert$

$= 10/15$

$d_{m,n;.025} = 7/15$ (PFA = .05)

$d_{m,n;.005} = 8/15$ (PFA=.01)

[Critical values are from Conovor (1971, p. 399)

Decision Rule:

Decide N + S since

$D_{m,n} > d_{m,n;\alpha/2}$ (α=.01).

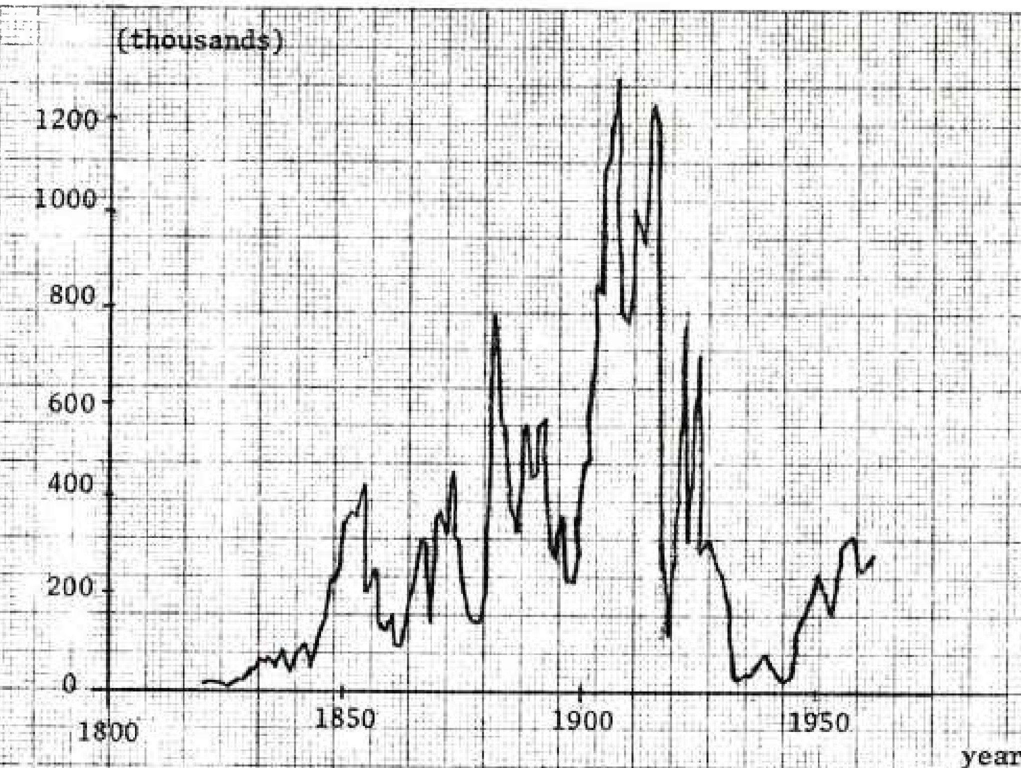


Fig. (a)

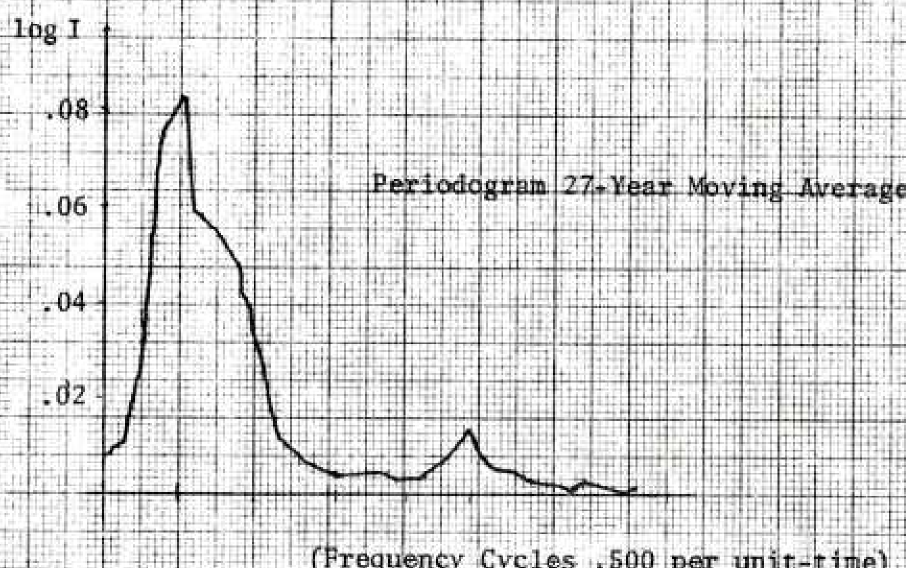
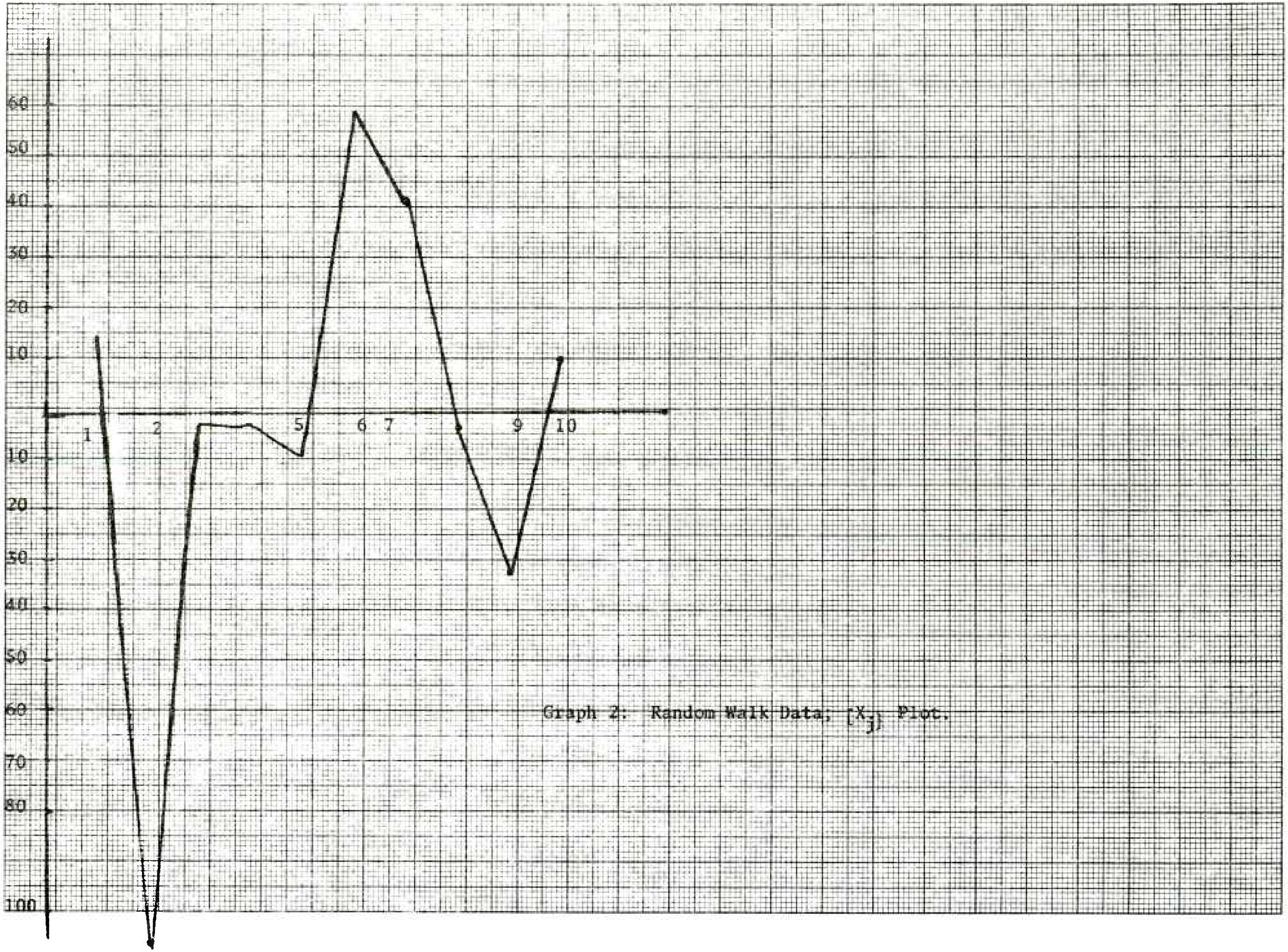
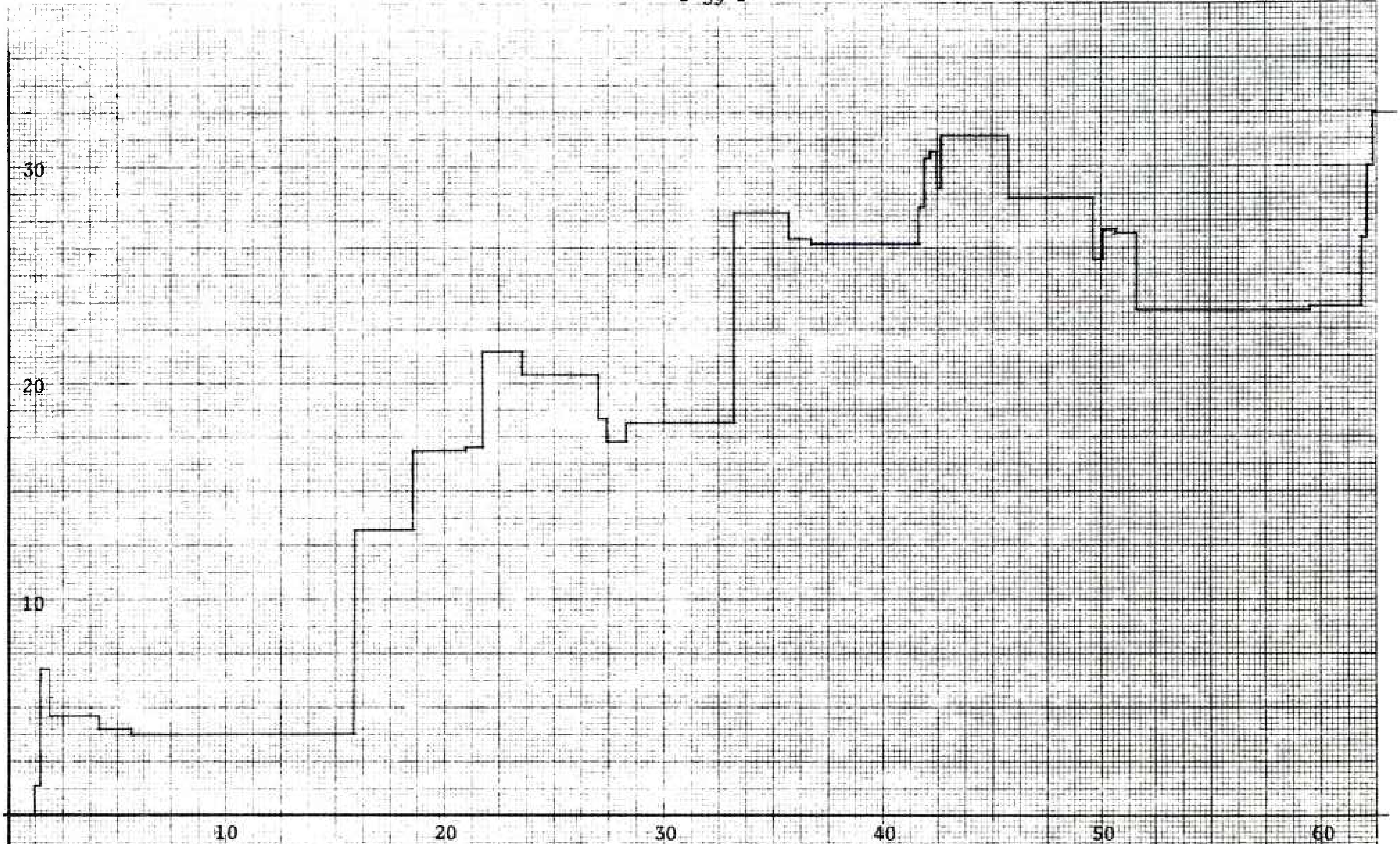


Fig. (b)

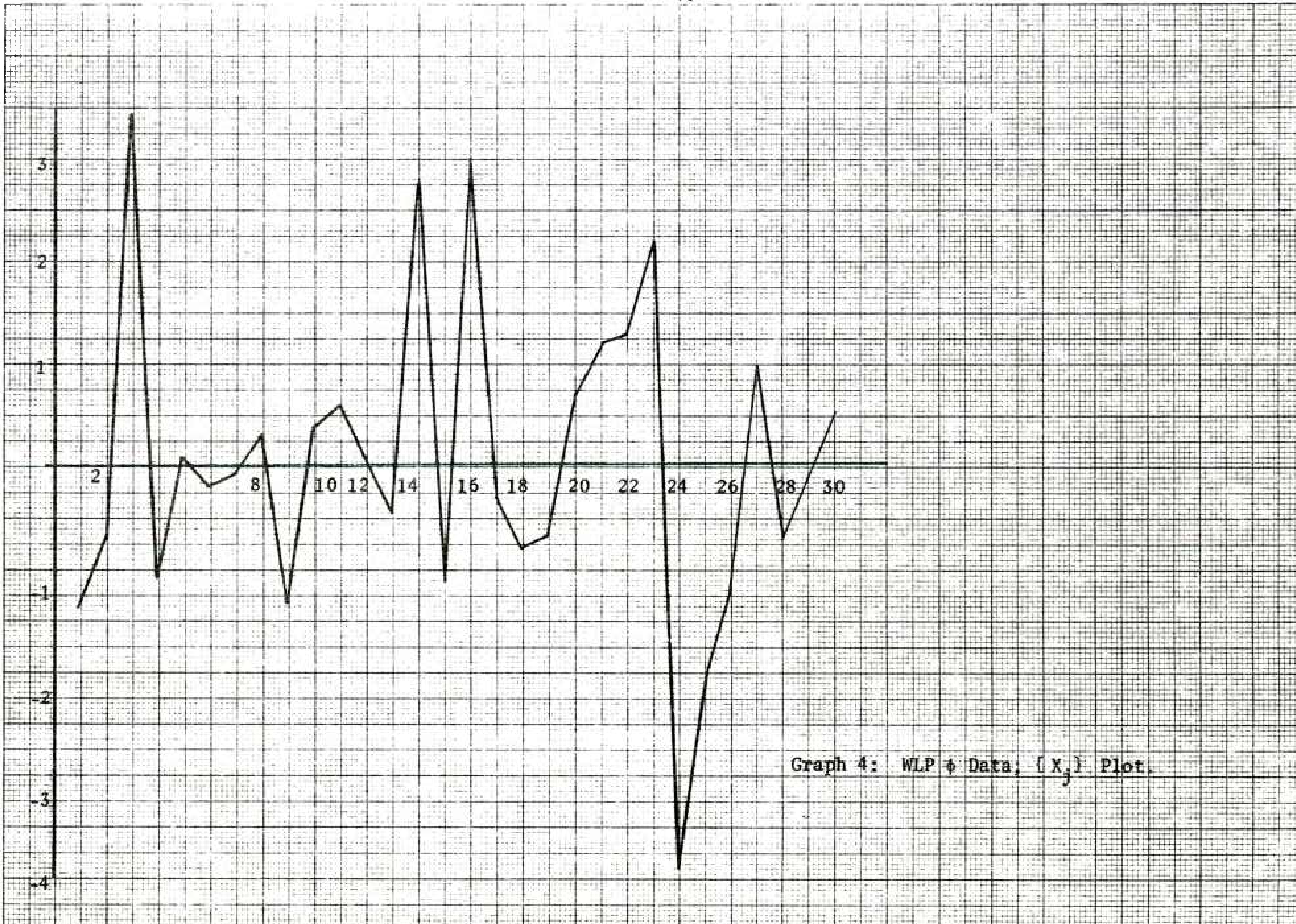
Graph 1: Immigration Data; Dewey (1963)



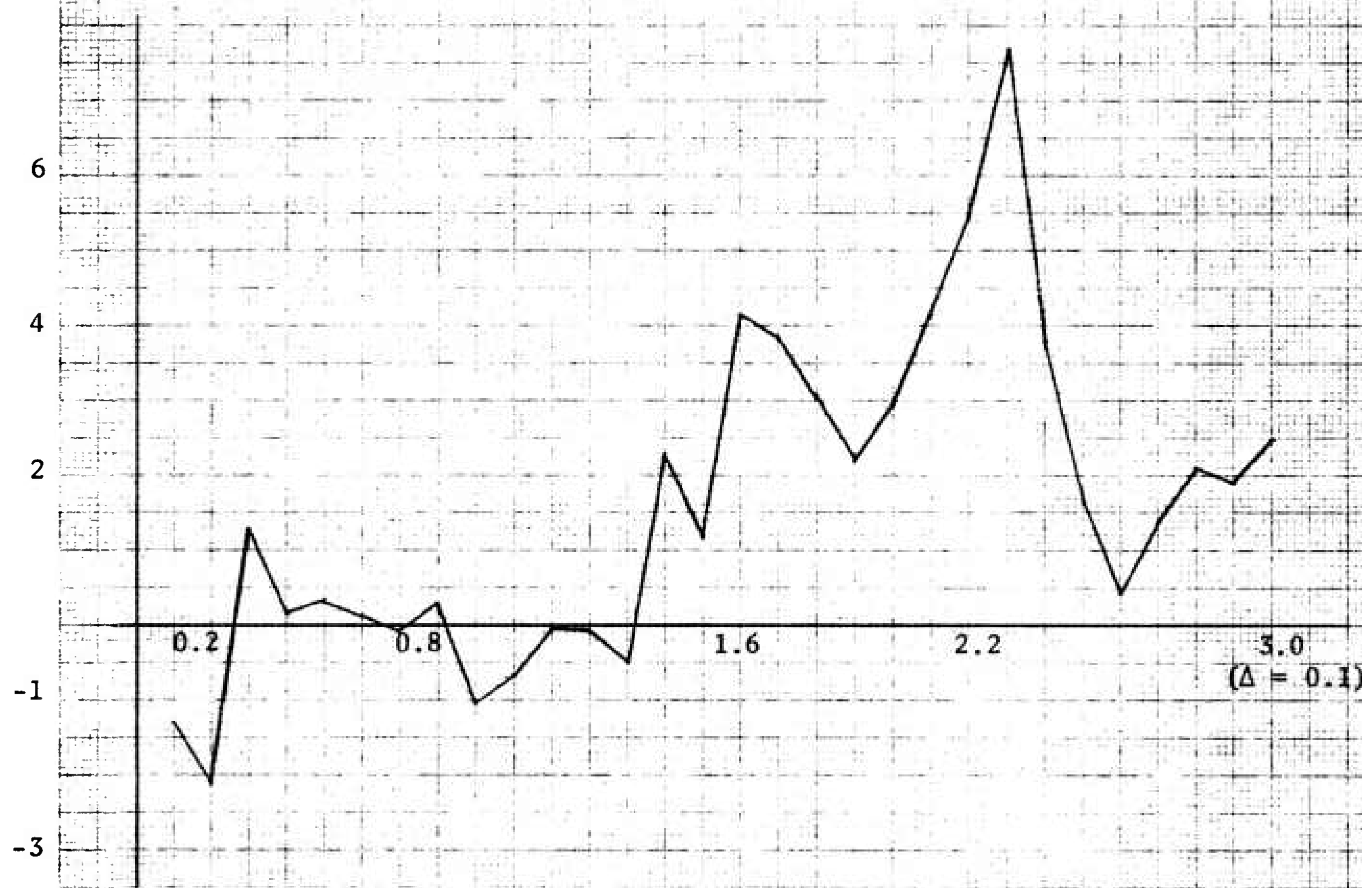
Graph 2: Random Walk Data; $\{X_j\}$ Plot.



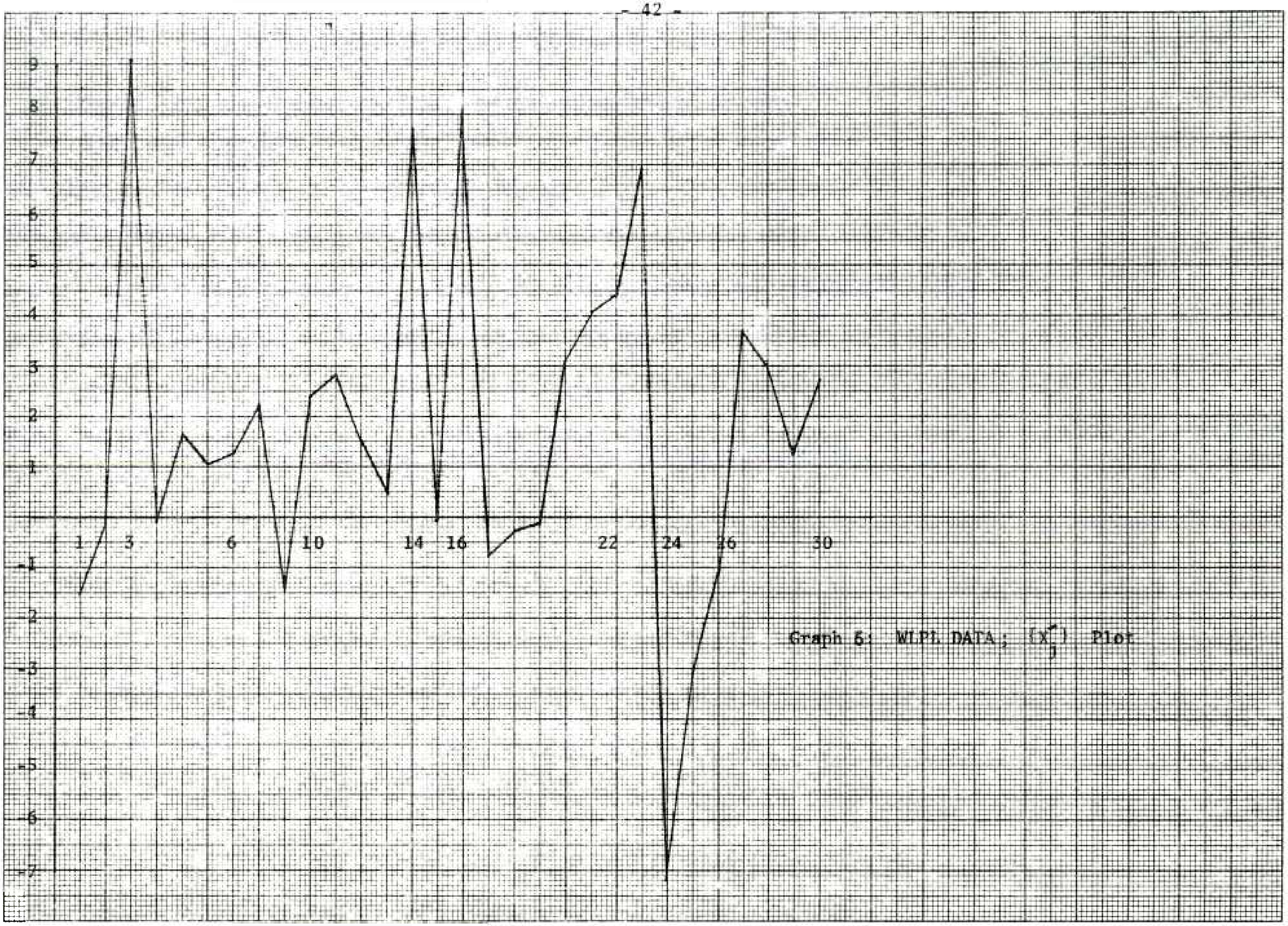
Graph 3: Compound Poisson Process; Plot for Jumps



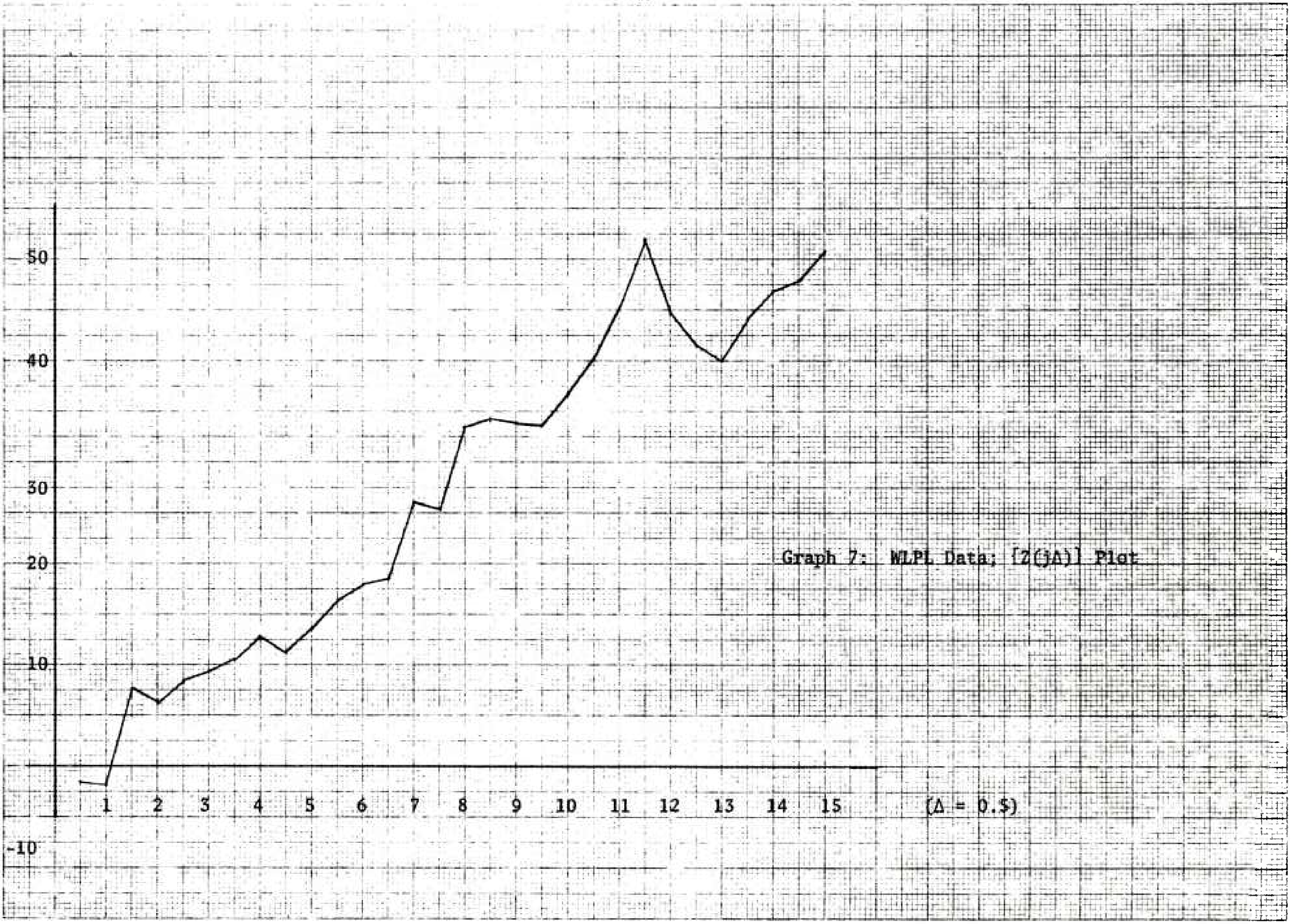
Graph 4: WLP ϕ Data; $\{X_j\}$ Plot.



Graph 5: WLP ϕ Data ; $\{Z(j\Delta)\}$ Plot.



Graph 6: WLPL DATA; $\{X_j^*\}$ Plot



Graph 7: WLPL Data; [Z(jΔ)] Plot